

# Detección de grupos de fajillas en imágenes de paquetes de billete en diversas condiciones de iluminación y fondo mediante un clasificador SVM

## Strips groups detection in images of banknote packages in different lighting and background conditions using an SVM classifier

Daniel Mauricio Florez Carvajal<sup>1</sup>, Germán Andrés Garnica Gaitán<sup>2</sup>

<sup>1</sup>Universidad Militar Nueva Granada, Bogotá, Colombia, dafo85@gmail.com

<sup>2</sup>Universidad Militar Nueva Granada, Bogotá, Colombia, german.garnica@unimilitar.edu.co

Fecha de recepción: 11/05/2017

Fecha de aceptación del artículo: 09/11/2017



Este obra está bajo una licencia de Creative Commons  
Reconocimiento-NoComercial-SinObraDerivada 4.0Internacional.

DOI: doi.org/10.18041/1794-4953/avances.11293

### Cómo citar:

Florez Carvajal, D., & Garnica Gaitán, G. (2017). Detección de grupos de fajillas en imágenes de paquetes de billete en diversas condiciones de iluminación y fondo mediante un clasificador SVM. AVANCES: Investigación en Ingeniería, 14, 145-154. doi.org/10.18041/1794-4953/avances.11293.

### Resumen

Este artículo presenta los resultados de una clasificación binaria de imágenes con dos diferentes condiciones de iluminación y fondo para un problema específico de detección de grupos de fajillas en paquetes de billete. La detección se lleva a cabo con un clasificador "Support Vector Machines" entrenado con vectores característicos obtenidos de las imágenes mediante la aplicación de la transformada wavelet y de la técnica de concatenación de histograma. Para cada condición de fondo e iluminación se entrena un clasificador diferente, se obtiene la matriz de confusión de cada uno y luego se comparan mediante los parámetros de recall, especificidad, precisión, exactitud y Fscore.

**Palabras claves:** Clasificación binaria, concatenación de histograma, matriz de confusión, support vector machines, transformada wavelet.

### Abstract

This article shows the results of a binary classification of images with two different lighting and background conditions for a specific case of strips groups detection in banknote packages. The detection is made with a support vector machine classifier trained with features obtained from the images through the application of the wavelet transform and histogram concatenation techniques. For each lighting and background condition a different classifier is trained, the confusion matrix is obtained for each one and then are compared through the recall, specificity, precision, accuracy and Fscore parameters.

**Keywords:** Binary classification, confusion matrix, histogram concatenation, support vector machines, wavelet transform.

## 1. Introducción

El recibo de consignaciones provenientes de bancos comerciales es uno de los procesos que debe llevar a cabo el Banco de la Republica de Colombia para cumplir con sus funciones encomendadas en la constitución [1]. Las consignaciones son realizadas en paquetes que deben contener cien fajos de billete agrupados en grupos de cinco, los cuales son inspeccionados y contados de manera visual por un trabajador de la entidad para dar conformidad a la consignación.

En ocasiones no se detectan inconsistencias en los paquetes consignados lo que conlleva consecuencias para el Banco de la República como descuadres contables y trámites administrativos, y consecuencias para el trabajador como presentación de informes, pago de diferencias y procesos disciplinarios. Esta situación motiva a proponer una solución tecnológica que apoye a los funcionarios en dicha labor.

Podría pensarse en soluciones triviales a la detección de fajos faltantes como pesar el paquete y compararlo con un valor de referencia, sin embargo, esto presentaría complicaciones ya que un billete muy deteriorado, por el polvo y la humedad que va adquiriendo a través de los años, presenta un grosor y un peso mayores a los de un billete en buen estado. Adicionalmente, el procedimiento de la entidad establece que la verificación debe realizarse de manera visual mediante conteo de grupos de fajillas [2].

La visión artificial busca imitar el proceso de percepción visual que los humanos realizamos y ha venido utilizándose en diferentes aplicaciones como reconocimiento óptico

de caracteres, inspección de maquinaria, ventas al por menor, imágenes médicas, seguridad automotriz, vigilancia, entre otros. Además, ha servido como un sistema de control de calidad en diferentes industrias [3]. Por esta razón se contempla como una opción adecuada para implementar la solución tecnológica al proceso.

El reconocimiento de grupos de fajillas en paquetes de billetes corresponde a un problema de clasificación binaria donde un algoritmo se encarga de establecer si una imagen pertenece o no a una categoría específica. Como casi todo problema de visión artificial requiere de una serie de pasos para lograr el objetivo como lo son un preprocesamiento de las imágenes para adecuarlas de forma que facilite el análisis posterior de éstas, una obtención de características del objeto a reconocer que lo diferencien de los demás y una clasificación basada en dichas características. La eficiencia de la clasificación depende de las características utilizadas para describir los objetos y de la capacidad del algoritmo para obtenerlas en diferentes condiciones de iluminación, fondo, ruido, rotación, entre otras variables que pueden afectar [4][5].

A continuación, se presenta un algoritmo de clasificación de grupos de fajillas en imágenes de paquetes de billete el cual es concebido como una fase inicial para el apoyo al proceso de inspección y conteo llevado a cabo por los funcionarios del Banco de la Republica de Colombia. Sirve como base para algoritmos adicionales que ejecuten las demás tareas realizadas en la inspección manual como lo son el conteo de fajillas y la verificación de códigos numéricos impresos en estos.

El algoritmo es probado en dos condiciones de iluminación y fondo diferentes. La primera corresponde a condiciones naturales del recinto, sin control de iluminación ni fondo que puedan generar brillos indeseados y ruido en la adquisición de las fotografías, y la segunda a condiciones controladas mediante un dispositivo de adquisición fabricado para controlar el fondo y disminuir los efectos de la iluminación del recinto sobre las imágenes.

A las imágenes de paquetes de billete se les aplica un preprocesamiento basado en la transformación del espacio de color original RGB, donde el color es definido mediante la adición de los colores rojo (R), verde (G) y azul (B), al espacio de color YCbCr, donde el color es definido en términos de una componente de luminancia (Y), una de crominancia azul (Cb) y otra de crominancia roja (Cr).

Posteriormente, a la componente de luminancia se le aplica una transformada wavelet. Sobre la imagen transformada se realiza un proceso de recorte y redimensionamiento, con lo que se obtienen imágenes de zonas que corresponden a grupos de fajillas y que no corresponden a grupos de fajillas.

A dichos recortes se les aplica la técnica de concatenación de histograma para conformar el conjunto de datos utilizados para el entrenamiento de un clasificador "Support Vector Machine – SVM" por cada condición de iluminación y fondo.

Los clasificadores son evaluados mediante una validación cruzada aleatoria de cuatro iteraciones de donde se obtiene la matriz de confusión de cada una, se promedian y luego se comparan mediante los parámetros de desempeño recall, especificidad, precisión, exactitud y Fscore.

## 2. Banco de imágenes

Para realizar la detección de grupos de fajillas se obtuvieron fotografías en dos condiciones de fondo y luz diferentes. La primera condición corresponde a fondo e iluminación naturales del recinto donde se lleva a cabo la actividad de recibo de paquetes, y la segunda a fondo e iluminación controlados mediante un dispositivo fabricado para tal fin.

En total se dispone de cien imágenes de paquetes, de las cuales cincuenta pertenecen a la primera condición y cincuenta a la segunda. En la Figura 1 se puede observar una imagen de cada condición trabajada.



**Figura 1.** Imágenes de paquetes de billete. (a) En Condiciones naturales y (b) Con fondo e iluminación controlados.

## 3. Preprocesamiento

Al banco de imágenes se le realiza un preprocesamiento para facilitar la obtención de información que permita diferenciar los grupos de fajillas y reducir la dimensionalidad de los datos, lo que disminuye los recursos computacionales y el tiempo de procesamiento necesario para la detección. Por lo tanto, se plantea la necesidad de representar la imagen a color original en una imagen de escala de grises en la que las fajillas puedan diferenciarse y tener características en común sin importar su color.

Del banco de imágenes se pudo determinar que las fajillas, al presentar una textura lisa, tienden a brillar más que los demás elementos, por lo que una representación en espacio de grises en la que se puedan diferenciar objetos brillantes es adecuada.

El espacio de color YCbCr, que define el color en términos de una componente de luminancia (Y), una de crominancia azul (Cb) y otra de crominancia roja (Cr), ofrece las características antes descritas en la componente de luminancia al ser independiente del color y basarse en el brillo [6].

Por lo anterior y por la reducción de tamaño de datos a la tercera parte de los originales, que surge por el hecho de trabajar con una sola componente, fue escogida la componente de luminancia para transformar el banco de imágenes a escala de grises. En la Figura 2 se aprecia la componente de luminancia del paquete mostrado en la Figura 1-(b).



**Figura 2.** Componente de luminancia del paquete mostrado en la figura 1-(b).

Con el fin de obtener una mejor diferenciación entre los grupos de fajillas y los demás objetos de la imagen se decide utilizar descriptores de textura debido a

la notable diferencia de ésta entre fajillas y demás elementos, lo que permite evidenciar a simple vista que los bordes de los grupos de fajillas están bien definidos sobre los billetes que envuelven como se aprecia en la Figura 2.

Para detectar los bordes se utiliza la transformada wavelet en dos dimensiones. Para el caso de una dimensión la transformada wavelet puede definirse como una descomposición de la señal en una familia de funciones obtenidas mediante el escalamiento y variación de frecuencia de una función base [7].

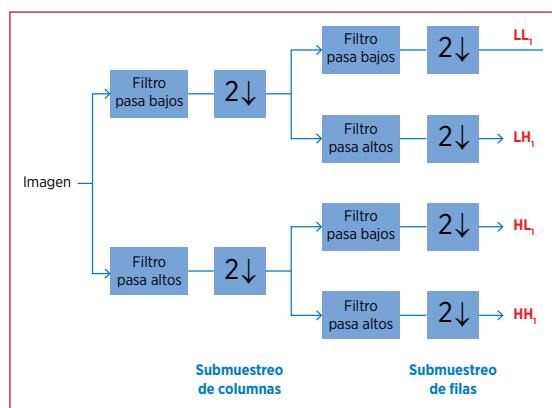
Esta descomposición para el caso discreto puede representarse mediante la aplicación de dos filtros espejo, uno de bajas frecuencias y otro de altas, de donde resultan dos señales con el doble de muestras de la original, por lo que se debe realizar un submuestreo a la mitad de cada una de ellas. Para el caso de dos dimensiones, como lo son las imágenes, la transformada wavelet puede verse como la aplicación de una transformada unidimensional discreta a lo largo de las filas y otra a lo largo de las columnas con lo que se obtiene una descomposición en cuatro bandas frecuenciales [8].

En la Figura 3 se observan los bancos de filtros utilizados para aplicar la transformada wavelet a una imagen. En esta se aprecia que de acuerdo con la serie de filtros utilizados se obtienen las cuatro bandas mencionadas anteriormente. Si se utiliza el filtro de bajas frecuencias tanto para las columnas como para las filas se obtiene la banda LL la cual mantiene los detalles de la imagen original y puede descomponerse nuevamente hasta el nivel deseado. Si se utiliza el filtro de bajas para las columnas y el de

altas para las filas se obtiene la banda LH con la que se pueden apreciar los bordes horizontales. Si se utiliza el filtro de altas para las columnas y el de bajas para las filas se obtiene la banda HL con la que se pueden apreciar los bordes verticales. Si se utiliza el filtro de altas frecuencias tanto para columnas como para filas se obtiene la banda HH donde se aprecian las esquinas.

Las bandas mostradas en la Figura 3 corresponden al primer nivel de profundidad. Si la banda LL se pasa por la misma serie de filtros se obtienen cuatro bandas adicionales que corresponden al segundo nivel de profundidad. Este proceso iterativo se puede seguir realizando de acuerdo al nivel de detalle y reducción de datos que se desee. La profundidad suele indicarse con un subíndice al a un costado de la banda respectiva.

De las cuatro bandas se escoge la HL para obtener los bordes verticales de los grupos de fajillas ya que ocupan una mayor área que los horizontales y no se ven afectados por los billetes, que al estar posicionados de manera horizontal tienden a generar ruido en la banda LH.



**Figura 3.** Bancos de filtros para aplicar la transformada de wavelet.

Se utiliza el nivel dos de la transformada debido a que se resaltan mejor los detalles de bordes, además de reducir el tamaño de los datos de la imagen de luminancia a la octava parte gracias al submuestreo que se realiza. Una transformada de nivel tres o superior, a pesar de reducir aún más los datos, tiende a distorsionar la imagen y por ende se pierden los detalles que permiten diferenciar los grupos de fajillas de los demás elementos.

En la Figura 4 se observan las cuatro bandas de nivel dos obtenidas al aplicar la transformada wavelet a una imagen de paquete de billetes. Allí se aprecia que en la banda HL los grupos de fajillas se distinguen de mejor manera que en las demás. Esta transformada se aplica a la componente de luminancia de todo el banco de imágenes.

La elección de esta transformada para caracterizar los grupos de fajillas obedece a los buenos resultados que presenta en problemas de clasificación de textura [7] [9][10] y a la superioridad que presenta en la clasificación respecto a otras técnicas [11].

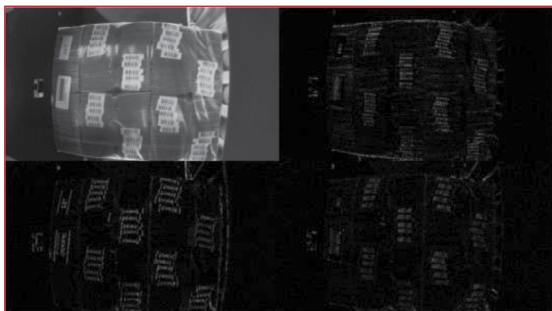
## 4. Conjunto de entrenamiento

Una vez se tiene la banda  $HL_2$  de todo el banco de imágenes, se procede a obtener el conjunto de datos para entrenar el clasificador SVM.

Para esto se realizan recortes de grupos de fajillas y de zonas que no corresponden a grupos de fajillas en la imagen de la banda  $HL_2$ . Los recortes se realizan de manera manual de acuerdo al tamaño de cada grupo de fajillas. Esta variabilidad de tamaño obedece a factores como la



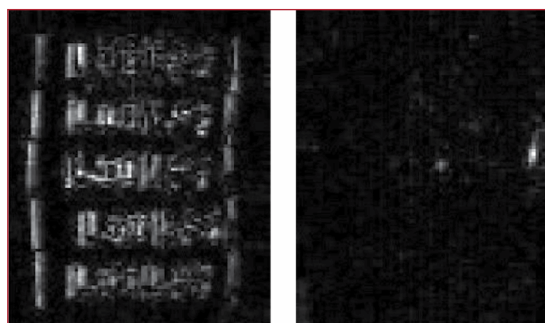
presión con que fueron ajustados los grupos de fajillas, la posición de las fajillas y el estado del billete.



**Figura 4.** Bandas obtenidas al aplicar la transformada wavelet de segundo nivel. Superior izquierda LL, superior derecha LH, inferior izquierda HL e inferior derecha HH.

Para poder entrenar el clasificador es necesario que los ejemplos de entrenamiento tengan la misma cantidad de datos característicos para cualquier ejemplo, por lo se realiza un redimensionamiento al mayor tamaño de recorte. Se utiliza el mayor debido a que la imagen resultante posee prácticamente la misma información que la original mientras que en un redimensionamiento a un tamaño menor, alguna información debe ser descartada abriendo la posibilidad de pérdida de detalles importantes para la caracterización [12].

En la Figura 5 se observan dos recortes redimensionados, uno de un grupo de fajillas y otro de una zona donde no hay grupos de fajillas.



(a)

(b)

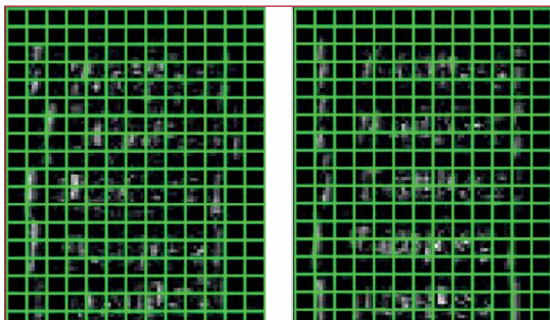
**Figura 5.** Recortes de la banda  $HL_2$  de (a) grupo de fajillas y (b) zona que no es grupo de fajillas.

De los recortes redimensionados se logró observar que los bordes de las fajillas presentan variaciones en su posición dentro de la ventana debido a la rotación, el ancho de la fajilla y desplazamiento del paquete al momento de adquirir la imagen. Esto significa que para una misma posición de píxel se puede llegar a tener valores muy altos si el borde está posicionado sobre éste o valores muy bajos si no lo está, lo que introduce error en el entrenamiento del clasificador.

Para reducir dicho error se utiliza la técnica de concatenación de histograma, la cual consiste en dividir la imagen en subregiones a las que se les obtiene el histograma para luego concatenarlos todos, de manera ordenada, en un vector al que se denomina vector característico. El hecho de dividir la imagen en subregiones permite un margen

de tolerancia para que los bordes de los grupos de fajillas puedan variar su posición sin alterar significativamente la distribución de píxeles obtenida por el histograma, ni la localización de dicha distribución en el vector característico.

En la Figura 6 se puede observar un ejemplo de dos grupos de fajillas cuyos bordes se encuentran en diferente posición, pero gracias a la división por subregiones son localizados en el mismo recuadro o en uno cercano, manteniendo un patrón similar de distribución y localización de intensidades de píxel en el vector característico de cada grupo de fajillas.



**Figura 6.** División por subregiones de la banda  $HL_2$  de dos grupos de fajillas con bordes en diferente posición.

Los vectores característicos de los recortes de grupos de fajillas conforman los ejemplos positivos y los vectores característicos de los recortes de zonas que no corresponden a grupos de fajillas conforman los ejemplos negativos. En total se dispone de 800 ejemplos para cada condición de iluminación, de los cuales 400 son positivos y 400 negativos.

## 5. Resultados y análisis

La evaluación de cada clasificador se realiza mediante la técnica de validación

cruzada aleatoria de cuatro iteraciones, la cual consiste en dividir los ejemplos en cuatro subconjuntos, cada uno de ellos con diferentes ejemplos de entrenamiento y prueba [13]. Con los subconjuntos se entrenan cuatro clasificadores diferentes, a los cuales se les obtiene la matriz de confusión y los parámetros de desempeño que finalmente son promediados para obtener un único resultado.

La matriz de confusión consiste en organizar las predicciones de un conjunto de datos de validación arrojadas por un modelo de aprendizaje supervisado en cuatro casillas [14]:

- **Verdadero Positivo (TP):** Ejemplos que el modelo predijo correctamente como positivos.
- **Falso Negativo (FN):** Ejemplos positivos que el modelo predijo incorrectamente como negativos.
- **Falso Positivo (FP):** Ejemplos negativos que el modelo predijo incorrectamente como positivos.
- **Verdadero Negativo (TN):** Ejemplos que el modelo predijo correctamente como negativos.

La suma de todas las casillas da como resultado el total de ejemplos de validación utilizados.

Los parámetros de desempeño se obtienen a partir de la matriz de confusión y son los siguientes [15]:

- **Sensitividad o Recall:** Es la proporción de ejemplos positivos que el modelo predijo correctamente como positivos y se calcula mediante la ecuación (1).

$$Recall = \frac{TP}{TP+FN} \quad (1)$$

- **Precisión:** Es la proporción de predicciones positivas que corresponden realmente a positivos y se calcula mediante la ecuación (2).

$$Precisión = \frac{TP}{TP+FP} \quad (2)$$

- **Especificidad:** Es la proporción de ejemplos negativos que el modelo predijo correctamente como negativos y se calcula mediante la ecuación (3).

$$Especificidad = \frac{TN}{FP+TN} \quad (3)$$

- **Exactitud:** Es la proporción de ejemplos que el modelo predijo correctamente respecto al total de ejemplos utilizados y se calcula mediante la ecuación (4).

$$Exactitud = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

- **Fscore:** Es un promedio ponderado de la precisión y la sensibilidad. Se calcula mediante la ecuación (5).

$$Fscore = \frac{2TP}{2TP+FP+FN} \quad (5)$$

En las Tablas 1 y 2 se muestran la matriz de confusión y los parámetros de desempeño resumidos para las dos condiciones de iluminación trabajadas.

**Tabla 1.** Resultados de evaluación del clasificador para imágenes en condiciones naturales.

Iteración	1	2	3	4	Promedio
<b>Recall</b>	1	0.9800	0.9700	1	0.9875
<b>Precisión</b>	1	0.9800	0.9238	0.9901	0.9735
<b>Especificidad</b>	1	0.9800	0.9200	0.9900	0.9725
<b>Exactitud</b>	1	0.9800	0.9450	0.9950	0.9800
<b>Fscore</b>	1	0.9800	0.9463	0.9950	0.9803
<b>VP</b>	100	98	97	100	NA
<b>VN</b>	100	98	92	99	NA
<b>FP</b>	0	2	8	1	NA
<b>FN</b>	0	2	3	0	NA

**Tabla 2.** Resultados de evaluación del clasificador para imágenes en condiciones controladas.

Iteración	1	2	3	4	Promedio
<b>Recall</b>	0.9500	0.9700	0.9700	0.9900	0.9700
<b>Precisión</b>	1	1	0.9898	0.9802	0.9925
<b>Especificidad</b>	1	1	0.9900	0.9800	0.9925
<b>Exactitud</b>	0.9750	0.9850	0.9800	0.9850	0.9813
<b>Fscore</b>	0.9744	0.9848	0.9798	0.9851	0.9810
<b>VP</b>	95	97	97	99	NA
<b>VN</b>	100	100	99	98	NA
<b>FP</b>	0	0	1	2	NA
<b>FN</b>	5	3	3	1	NA



De los resultados obtenidos se observa que para las dos condiciones de iluminación y fondo los clasificadores presentan resultados similares con una leve mejoría en el caso controlado. Esto demuestra que las técnicas utilizadas para el procesamiento de las imágenes y adquisición de características son robustas ante brillos y fondos complejos.

En ambos casos el algoritmo arroja un promedio de 98% de exactitud, lo que significa una muy alta proporción de ejemplos correctamente catalogados en las dos condiciones. Sin embargo, al revisar los ejemplos que fueron mal catalogados, se aprecia que en condiciones naturales el algoritmo presenta más falsos positivos que falsos negativos, y en condiciones controladas sucede lo contrario.

Lo anterior se ve reflejado en los parámetros recall, precisión y especificidad. El recall es 1.75% mayor en condiciones naturales, es decir que en estas condiciones el clasificador tiende a cometer menos el error de catalogar como “diferente a grupo de fajillas” una imagen que en realidad si corresponde a un grupo de fajillas. Los parámetros precisión y especificidad son 1.9% y 2% respectivamente superiores en condiciones controladas, es decir que en estas condiciones el clasificador tiende a cometer menos el error de catalogar una imagen que no corresponde a un grupo de fajillas como si así lo fuera.

Esta diferencia es importante ya que, a pesar de que es necesario detectar paquetes de billete que contengan más o menos fajos de los establecidos por las directrices, si un paquete con menos fajos no se detecta y es aceptado por los funcionarios que realizan la inspección, las consecuencias son más severas que si el paquete trae más

fajos porque el funcionario deberá asumir el pago del dinero faltante y enfrentar un proceso disciplinario.

Por lo tanto, a pesar de que el algoritmo presenta muy buenos resultados en las dos condiciones probadas, se prefiere la adquisición de imágenes en un ambiente controlado.

## 6. Conclusiones

El análisis del banco de imágenes fue de vital importancia ya que permitió establecer las variables críticas a tener en cuenta y escoger las técnicas adecuadas para su tratamiento.

Las características con las que se describen los objetos a buscar y las técnicas para obtenerlas son determinantes en la efectividad del clasificador. Es importante que las características escogidas sean un patrón repetitivo que permita al clasificador un correcto aprendizaje.

Los altos porcentajes de desempeño del algoritmo, tanto en condiciones naturales como controladas, indican su robustez ante variables externas como brillos, iluminación y fondo.

Los resultados mostrados afirman que el algoritmo es efectivo en las dos condiciones probadas y por ende puede ser utilizado como fase de segmentación en un sistema que realice la inspección de todas las variables que deben tenerse en cuenta para dar conformidad a una consignación.

Se prefiere la adquisición de imágenes mediante el dispositivo que controla la iluminación y el fondo ya que el algoritmo tiene menos probabilidad de cometer falsos

positivos. Además, se controla la aparición de nuevas condiciones del ambiente de operación como consecuencia de cambio de luminarias o cambio de lugar de operación.

## 7. Agradecimientos

Los autores de este trabajo quieren agradecer al Grupo de Investigación ACCEDER de la Universidad Militar Nueva Granada, en especial a la profesora Marcela Iregui. Este trabajo fue financiado por la Vicerrectoría de Investigaciones de la Universidad Militar Nueva Granada, en el marco de proyectos de investigación científica con el proyecto INV ING 2097, vigencia 2016.

## 8. Referencias

- [1] Imprenta Nacional (1991). Constitución política de Colombia. Santafé de Bogotá.
- [2] Colombia, B. d. (2013). Manual de Tesorería. Circular Reglamentaria Externa - DTE - 164. Asunto 01: Consignaciones sujetas a verificación y operaciones de cambio de efectivo. Bogotá D.C.
- [3] Szeliski, R. (2010). Computer vision: algorithms and applications. Springer Science & Business Media.
- [4] Lu, D., Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International journal of Remote sensing*, 28, 5, 823-870.
- [5] Athinarayanan, S., Srinath, M., Kavitha, R. (2016). Detection and Classification of Cervical Cancer in Pap Smear Images using EETCM, EEETCM & CFE methods based Texture features and Various Classification Techniques. *IJSRSET*, 2, 5.
- [6] Lin, C. (2007). Face detection in complicated backgrounds and different illumination conditions by using YCbCr color space and neural network. *Pattern Recognition Letters*, 28, 16, 2190-2200.
- [7] Ahmadian, A., Mostafa, A. (2003). An efficient texture classification algorithm using Gabor wavelet. *Engineering in Medicine and Biology Society. Proceedings of the 25th Annual International Conference of the IEEE*, 1, 930-933.
- [8] Mallat, S. G. (1989). A Theory of Multiresolution Signal Decomposition - The Wavelet Representation. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 2, 7, 674-693.
- [9] Chang, T., Kuo, C.-C. (1993). Texture analysis and classification with tree-structured wavelet transform. *IEEE Transactions on image processing*, 2, 4, 429-441.
- [10] Ghazvini, M., Monadjemi, S. A., Movahhedinia, N., Jamshidi, K. (2009). Defect detection of tiles using 2D-wavelet transform and statistical features. *World Academy of Science, Engineering and Technology*, 49, 901-904.
- [11] Manjunath, B. S., Ma, W. (1996). Texture features for browsing and retrieval of image data. *IEEE Transactions on pattern analysis and machine intelligence*, 18, 8, 837-842.
- [12] Dollár, P., Belongie, S., Perona, P. (2010). The Fastest Pedestrian Detector in the West. *BMVC*, 2, 3, 7.
- [13] Arlot, S., Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys*, 4, 40-79.
- [14] Davis, J., Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. *Proceedings of the 23rd international conference on Machine learning*, 233-240.
- [15] Powers, D. M. (2007). Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation, School of Informatics and Engineering, *Journal of Machine Learning Technologies*, 2, 1, 37-63.